Molecular Cloning

A LABORATORY MANUAL

SECOND EDITION

J. Sambrook
UNIVERSITY OF TEXAS SOUTHWESTERN MEDICAL CENTER

E.F. Fritsch
GENETICS INSTITUTE

T. Maniatis
HARVARD UNIVERSITY

Cold Spring Harbor Laboratory Press
1989

Bacteriophage SP6 and Bacteriophages T7 and T3 RNA Polymerases

(Bacteriophage SP6-infected *Salmonella typhimurium* LT2 and bacteriophage T7- or T3-infected *E. coli*)

Bacteriophage SP6 synthesizes a DNA-dependent RNA polymerase that recognizes and initiates synthesis of RNA on double-stranded DNA templates that carry the appropriate bacteriophage-specific promoter. The polymerase is used in vitro to generate large quantities of RNA complementary to one strand of foreign DNA that has been cloned immediately downstream from the promoter in plasmids specifically designed for this purpose. Vectors are available to synthesize RNA complementary to either strand of the template by changing the orientation of the promoter with respect to the cloned foreign DNA sequences (Butler and Chamberlin 1982; Melton et al. 1984).

Bacteriophages T7 and T3 also synthesize DNA-dependent RNA polymerases that recognize and initiate synthesis of RNA on double-stranded DNA templates that carry the appropriate bacteriophage-specific promoter. These polymerases are used in vitro just like the bacteriophage SP6 RNA polymerase. Bacteriophages T7 and T3 RNA polymerases have been cloned and expressed in *E. coli* (Davanloo et al. 1984; Tabor and Richardson 1985; Morris et al. 1986) and bacteriophage T7 RNA polymerase has been cloned and expressed in yeast (Chen et al. 1987). This allows vectors carrying the bacteriophage T7 promoter to be used to express cloned genes in vivo.

USES

1. Synthesis of single-stranded RNA for use as hybridization probes, functional mRNAs for in vitro translation systems, or substrates for in vitro splicing reactions. Each of the three RNA polymerases has a high degree of specificity for its cognate promoter.

2. The bacteriophage T7 transcription system has been used to express cloned genes in bacteria (Tabor and Richardson 1985; Studier and Moffatt 1986) and in yeast (Chen et al. 1987).

   Two types of bacteriophage T7 expression systems have been developed for *E. coli*. In the first system, stable lysogens are established with bacteriophage λ carrying the bacteriophage T7 RNA polymerase gene under the control of the *E. coli lac*UV5 promoter. Plasmids containing the gene of interest under the control of the bacteriophage T7 promoter are then introduced into the lysogens containing the bacteriophage T7 RNA polymerase gene. Activation of the bacteriophage T7 promoter is then achieved by isopropylthio-β-D-galactoside induction of the *lac*UV5 promoter driving the bacteriophage T7 RNA polymerase gene. In the second system, the bacteriophage T7 promoter/plasmid carrying the gene of interest is introduced into bacteria, and the bacteriophage T7 promoter is activated by infecting the bacteria with bacteriophage λ containing the bacteriophage T7 RNA polymerase gene.

   In yeast, the bacteriophage T7 RNA polymerase gene is placed under the control of a yeast promoter and stably introduced into yeast cells on an autonomously replicating vector. Expression is achieved by introducing into the yeast cells a second plasmid that contains the gene of interest under the control of the bacteriophage T7 promoter (Chen et al. 1987).

FUNCTIONAL COMPONENTS OF MAMMALIAN EXPRESSION VECTORS

Mammalian expression vectors contain both prokaryotic sequences that facilitate the propagation of the vector in bacteria and one or more eukaryotic transcription units that are expressed only in eukaryotic cells. The eukaryotic transcription unit consists of noncoding sequences and sequences coding for selectable markers. It is frequently assembled as a composite of elements derived from different, well-characterized viral or mammalian genes. The components that are used in various expression vectors are described briefly below.

Prokaryotic Plasmid Sequences That Facilitate the Construction, Propagation, and Amplification of Recombinant Vector Sequences in Bacteria

The essential prokaryotic elements include a replicon that functions in *Escherichia coli*, a gene encoding antibiotic resistance to permit selection of bacteria that harbor the recombinant plasmids, and a limited number of unique restriction sites in nonessential regions of the plasmid to allow insertion of eukaryotic sequences. Most of the mammalian vectors in current use contain prokaryotic sequences from derivatives of the plasmid pBR322 (e.g., pXf3, pBRd, and pML) that lack sequences that seem to interfere with the replication of the transfected DNA in eukaryotic cells (Lusky and Botchan 1981). Deletion of unnecessary segments of plasmid DNA also reduces the size of the vector and facilitates the placing of unique restriction sites that can be utilized for the insertion and manipulation of eukaryotic sequences.

A Eukaryotic Expression Module That Contains All of the Elements Required for the Expression of Foreign DNA Sequences in Eukaryotic Cells

The most basic eukaryotic expression module contains a promoter element to mediate transcription of foreign DNA sequences and signals required for efficient polyadenylation of the transcript. Additional elements of the module may include enhancers and introns with functional splice donor and acceptor sites.

PROMOTER AND ENHANCER ELEMENTS

Unlike the signals required for RNA processing, which function efficiently in all types of mammalian cells, the activities of elements that control transcription—promoters and enhancers—vary considerably among different cell types. Promoters and enhancers consist of short arrays of DNA sequences that interact specifically with cellular proteins involved in transcription (for review, see Dynan and Tjian 1985; Serfling et al. 1985; McKnight and Tjian 1986; Sassone-Corsi and Borrelli 1986; Maniatis et al. 1987). The combination of different recognition sequences and the amounts of the cognate transcription factors determine the efficiency with which a given gene is transcribed in a particular cell type.

Many eukaryotic promoters contain two types of recognition sequences: the

TATA box and the *upstream promoter elements*. The TATA box, located 25–30 bp upstream of the transcription initiation site, is thought to be involved in directing RNA polymerase II to begin RNA synthesis at the correct site. In contrast, the upstream promoter elements determine the rate at which transcription is initiated. These elements can act regardless of their orientation, but they must be located within 100 to 200 bp upstream of the TATA box. *Enhancer elements* can stimulate transcription up to 1000-fold from linked homologous or heterologous promoters. However, unlike upstream promoter elements, enhancers are active when placed downstream from the transcription initiation site or at considerable distances from the promoter. Many enhancers of cellular genes work exclusively in a particular tissue or cell type (for review, see Voss et al. 1986; Maniatis et al. 1987). In addition, some enhancers become active only under specific conditions that are generated by the presence of an inducer, such as a hormone or metal ion (for review, see Sassone-Corsi and Borrelli 1986; Maniatis et al. 1987). Because of these differences in the specificities of cellular enhancers, the choice of promoter and enhancer elements to be incorporated into a eukaryotic expression vector will be determined by the cell type(s) in which the recombinant gene is to be expressed. Conversely, the use of a prefabricated vector containing a specific promoter and cellular enhancer may severely limit the cell types in which expression can be obtained.

Many enhancer elements derived from viruses have a broader host range and are active in a variety of tissues, although significant quantitative differences are observed among different cell types. For example, the SV40 early gene enhancer is promiscuously active in many cell types derived from a variety of mammalian species, and vectors incorporating this enhancer have consequently been widely used (Dijkema et al. 1985). Two other enhancer/promoter combinations that are active in a broad range of cells are derived from the long terminal repeat (LTR) of the Rous sarcoma virus genome (Gorman et al. 1982b) and from human cytomegalovirus (Boshart et al. 1985).

TERMINATION AND POLYADENYLATION SIGNALS

During the expression of eukaryotic genes, RNA polymerase II transcribes through the site where polyadenylation will occur. Consequently, the 3′ terminus of the mature mRNA is formed by site-specific posttranscriptional cleavage and polyadenylation (for review, see Birnstiel et al. 1985; Proudfoot and Whitelaw 1988; Proudfoot 1989). Although discrete sites for the termination of the primary transcript have not yet been identified, general regions of DNA a few hundred nucleotides in length and downstream from the polyadenylation site have been identified where transcription randomly terminates.

Two distinct sequence elements are required for accurate and efficient polyadenylation: (1) GU- or U-rich sequences located downstream from the polyadenylation site and (2) a highly conserved sequence of six nucleotides, AAUAAA, located 11–30 nucleotides upstream, which is necessary but not sufficient for posttranscriptional cleavage and polyadenylation (for review, see Mason et al. 1986; Proudfoot and Whitelaw 1988). The practical implication of these observations is that sequences downstream from the polyad

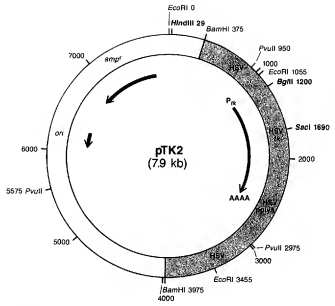

FIGURE 16.1A

pTK2 is a derivative of pBR322 that carries a 3.6-kb *Bam*HI fragment of herpes simplex virus (HSV) encoding thymidine kinase (*tk*). The positions of the *tk* promoter (P_{tk}) and the polyadenylation site (polyA; AAAA) are indicated.

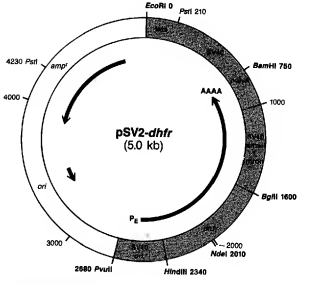

FIGURE 16.1B

pSV2-*dhfr* carries the SV40 origin (SV40 *ori*) and expresses dihydrofolate reductase (*dhfr*) from the SV40 early promoter (P_E). The SV40 small T intron and polyadenylation site (polyA; AAAA) are shown.

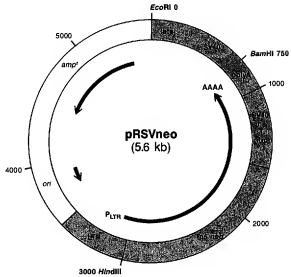

FIGURE 16.1C
pRSVneo expresses aminoglycoside phosphotransferase (APH) encoded by the bacterial transposon gene Tn*5* *neo*^{r} from the Rous sarcoma virus (RSV) LTR promoter (P_{LTR}). The SV40 small T intron and polyadenylation site (polyA; AAAA) are located downstream from Tn*5* *neo*.

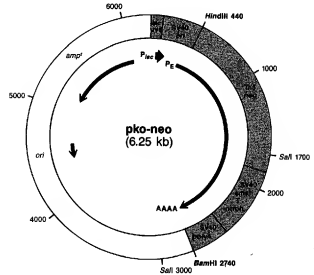

FIGURE 16.1D
pko-neo expresses aminoglycoside phosphotransferase encoded by the bacterial transposon gene Tn*5* *neo*^{r} from the eukaryotic SV40 early promoter (P_E) or the prokaryotic *E. coli lac*UV5 promoter (P_{neo}). The SV40 origin (SV40 *ori*), SV40 small T intron, and SV40 polyadenylation sites (polyA; AAAA) are present.

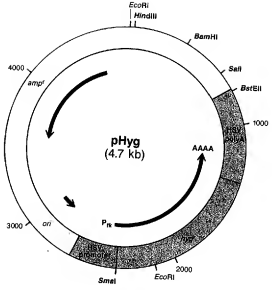

FIGURE 16.1E
pHyg directs the expression of the *E. coli* gene encoding hygromycin B phosphotransferase (*hyg*^{r}) using the herpes simplex virus promoter (P_R) and polyadenylation site (HSV polyA; AAAA).

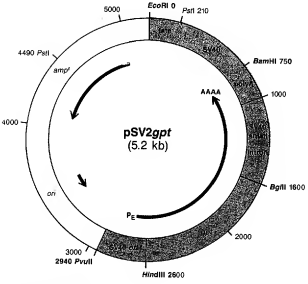

FIGURE 16.1F
In pSV2*gpt*, the *E. coli* xanthine-guanine phosphoribosyl transferase gene (*gpt*) is expressed using the SV40 early promoter (P_E) located in the SV40 origin (SV40 *ori*), the SV40 small T intron, and the SV40 polyadenylation site (polyA; AAAA).

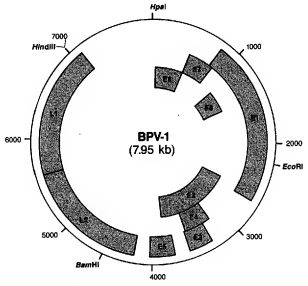

FIGURE 16.4
Bovine papillomavirus (BPV-1) encodes eight early gene products (E1–E8) and two late gene products (L1 and L2).

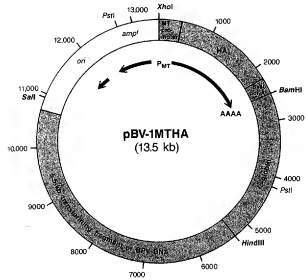

FIGURE 16.5
pBV-1MTHA carries the 69% transforming fragment (5.5-kb transforming fragment of BPV DNA), which allows episomal replication of this vector in mammalian cells. In addition, it encodes hemagglutinin (HA) expressed from the metallothionein (M_T) promoter (P_{MT}) and human β-globin. It also has an SV40 polyadenylation site (polyA; AAAA).

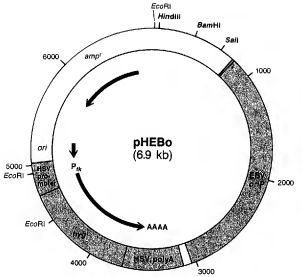

FIGURE 16.6A

pHEBo contains the Epstein-Barr virus origin *P* (EBV *oriP*) and directs the expression of the *E. coli* gene encoding hygromycin B phosphotransferase (*hyg*^{r}) using the herpes simplex virus promoter (P_{tk}) and polyadenylation site (HSV polyA; AAAA).

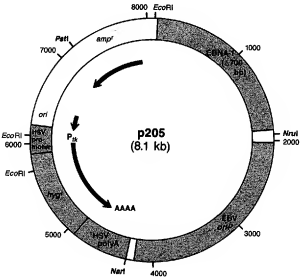

FIGURE 16.6B

p205 carries the Epstein-Barr virus origin *P* (EBV *oriP*) and expresses a *trans*-acting Epstein-Barr nuclear antigen (EBNA-1) from DNA containing a 700-bp deletion (Δ700 bp). It also directs the expression of the *E. coli* gene encoding hygromycin B phosphotransferase (*hyg*^{r}) using the herpes simplex virus promoter (P_{tk}) and polyadenylation site (HSV polyA; AAAA).

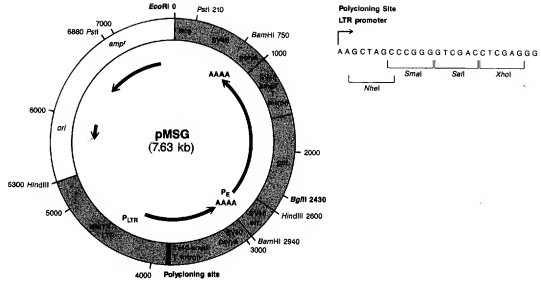

FIGURE 16.3A

pMSG is a transient expression vector that uses the mouse mammary tumor virus (MMTV) LTR promoter (P_{LTR}) to express cDNA sequences cloned into the polycloning site. The *E. coli* xanthine-guanine phosphoribosyl transferase gene (*gpt*) is expressed from the SV40 promoter (P_E) located in the SV40 origin (SV40 *ori*). The SV40 small T intron and polyadenylation site (SV40 polyA; AAAA) are found downstream from both the polycloning site and the *gpt* gene.

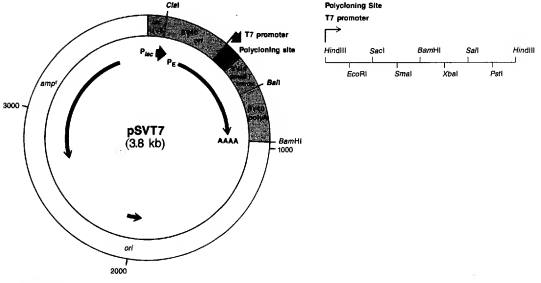

FIGURE 16.3B

pSVT7 is a transient expression vector that uses SV40 signals, the SV40 early promoter (P_E), the SV40 small T intron, and the SV40 polyadenylation site (polyA; AAAA) to express cDNA cloned into the polycloning site. The vector also carries the *lac*UV5 promoter (P_{lac}) for expression in *E. coli* and a bacteriophage T7 promoter for generation of mRNA in vitro.

and a 3′ splice site from a mouse immunoglobulin gene, (4) the SV40 polyadenylation signal, and (5) the adenovirus VA_I RNA gene region. pMT2 also carries sequences encoding murine DHFR positioned downstream from the splice acceptor site. cDNAs inserted between the splice acceptor site and the *dhfr* sequences can be transiently expressed at high levels in COS cells. The inserted cDNA is transcribed to produce a hybrid polycistronic mRNA in which the sequence coding for the foreign protein is flanked by the adenovirus tripartite leader and the murine *dhfr*. Because it lies at the 3′ terminus of the transcription unit, *dhfr* is inefficiently translated, but it can nevertheless serve as a selective, amplifiable marker (see page 16.28) and may also enhance the stability of the polycistronic mRNA (Kaufman et al. 1987 and unpubl.). The adenovirus tripartite leader and the VA_I RNA increase the efficiency of translation of the foreign coding sequences (3- to 20-fold) (Kaufman et al. 1985; Kaufman and Murtha 1987) by blocking the activity of a double-stranded RNA-dependent protein kinase that phosphorylates the α subunit of eukaryotic initiation factor 2 (eIF-2) (Kitajewski et al. 1986; O'Malley et al. 1986; Akusjärvi et al. 1987). (*Note:* Expression of the adenovirus VA RNAs is restricted to cells [e.g., COS cells] that contain a specific factor required for transcription of these RNA-polymerase-III-dependent genes.)

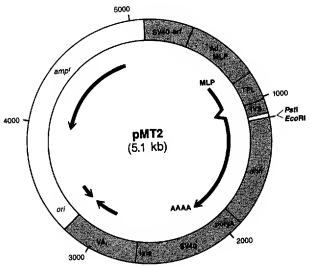

FIGURE 16.3C
pMT2 is a transient expression vector in which the adenovirus major late promoter (Ad MLP; MLP) is coupled to the adenovirus tripartite leader (TPL), which contains a 5′ splice site and part of a mouse immunoglobulin gene that contains a 3′ splice site. The intervening sequence (IVS) is followed by two cloning sites (*Pst*I and *Eco*RI), the dihydrofolate reductase gene (*dhfr*), the SV40 polyadenylation site (polyA; AAAA) and the adenovirus VA_I gene. The SV40 origin (SV40 *ori*) and enhancer are also present on the vector.

analysis of mutant promoters/enhancers with reporter genes should be confirmed by direct assay of the mRNAs encoded by critical mutants.

VECTORS CARRYING REPORTER GENES

A number of different prokaryotic genes have been used as reporters of the transcriptional activity of mammalian promoters. However, not all of these are suitable for use in transient expression systems. For example, the bacterial *gpt* (Mulligan and Berg 1980), *neo*^{r} (Schöler and Gruss 1984), and *gal*K (galactokinase) (Schümperli et al. 1982) genes can be expressed in mammalian cells when linked to a mammalian promoter. However, the assays for all three of these enzymes are cumbersome, requiring, for example, the use of starch gel electrophoresis to separate the bacterial enzymatic activity from activities endogenous to mammalian cells.

The following considerations are important when selecting a reporter gene (Gorman et al. 1982a):

- The enzymatic activity encoded by the prokaryotic gene must be readily distinguishable from any similar activities present in the mammalian cells prior to transfection
- There should be no interference or competition from other enzymatic activities in the cells
- The assay for the encoded enzymatic activity should be rapid, sensitive, reproducible, and convenient

The enzyme CAT fulfills all of these criteria and its gene has therefore become the most widely used reporter gene for indirect assay of promoter activity in transfected mammalian cells. The *cat* gene was originally derived from the transposable element Tn*9* and confers resistance to chloramphenicol. The coding region is 1102 bp in length, which is ordinarily flanked by two 768-bp IS elements. A plasmid, pSV2CAT, has been constructed that contains the SV40 promoter/enhancer, 29 bp of 5′ untranslated sequence, the CAT coding sequence, and 8 bp of DNA 3′ to the UAA stop codon. pSV2CAT cannot confer chloramphenicol resistance on bacteria because the *cat* gene is not linked to a prokaryotic promoter. To assay putative promoters in mammalian cells, a derivative of pSV2CAT has been constructed (pSV0CAT; Gorman et al. 1982a) in which the promoter region of SV40 is replaced by the promoter being tested. CAT modifies and inactivates chloramphenicol by mono- and diacetylation, and a number of convenient assays have been developed to measure CAT activity in mammalian cells transfected with derivatives of pSV0CAT, including:

- *Incubation of extracts prepared from transfected cells with ^{14}C-labeled chloramphenicol.* The extent of modification of chloramphenicol is measured by thin-layer chromatography on silica gels (see pages 16.60–16.62), which separates the mono- and diacetylated derivatives of chloramphenicol from the unmodified compound. The silica gel is exposed to X-ray film and aligned with the resulting autoradiograph. Regions of the gel corresponding to spots on the film are scraped from the plate, and the amount of radioactivity is measured in a liquid scintillation counter. This assay can be somewhat tedious when large numbers of samples are assayed.

- *Incubation of extracts prepared from transfected cells with unlabeled chloramphenicol and ^{14}C-labeled acetyl coenzyme A.* CAT catalyzes the transfer of the ^{14}C-labeled acetyl group from acetyl coenzyme A to chloramphenicol. At the end of the reaction, the mixture is extracted with ethyl acetate. Acetylated forms of chloramphenicol partition into the organic phase, whereas the acetyl coenzyme A remains in the aqueous phase. The amount of chloramphenicol converted to the acetylated form can then be measured in a liquid scintillation counter. Before using this assay to measure CAT activity, it is necessary to heat the cell extracts for 10 minutes at 65°C to destroy an activity that consumes acetyl coenzyme A. This problem is not encountered with the silica gel assay because large amounts of unlabeled acetyl coenzyme A are present in the reaction.

When measuring the effect of promoters and enhancers on gene expression, it is essential to include an internal control that will distinguish differences in the level of transcription from differences in the efficiency of transfection or in the preparation of extracts. This is best achieved by cotransfecting the cells with two plasmids—one that carries the construct under investigation and another that constitutively expresses an activity that can be assayed in the same cell extracts prepared for measurement of CAT activity. An enzyme frequently used for this purpose is *E. coli* β-galactosidase, which is expressed in transfected mammalian cells from a promoter with a broad host range (the SV40 early promoter or the Rous sarcoma virus LTR). Extracts of most types of mammalian cells contain relatively low levels of endogenous β-galactosidase activity, and an increase in enzyme activity of up to 100-fold can usually be detected during the course of a transfection. However, β-galactosidase should not be used as an internal control in certain specialized cells (e.g., gut epithelial cells) that express high levels of this activity.

A number of different approaches can be used to normalize the CAT activity to the β-galactosidase activity. For example, the amount of protein in individual extracts prepared from a series of transfected cells is measured, and the CAT and β-galactosidase assays (see below) are then carried out using a standard amount of protein in each assay. The CAT activity is then normalized to the β-galactosidase activity. Alternatively, the β-galactosidase activity in a constant volume of extract is measured and the CAT assay is then carried out with amounts of extract that contain a defined amount of β-galactosidase activity. Finally, both enzymatic assays can be carried out in a constant volume of extract, and the results can then be normalized to a defined level of β-galactosidase activity.

The human growth hormone gene has also been used as an internal control for transfection (Selden et al. 1986). In this case, the coding region of the growth hormone gene is linked to a metallothionein or to the Rous sarcoma virus LTR—promoters that are expressed in most types of cultured mammalian cells. Rather than assaying extracts of transfected cells, the amount of growth hormone secreted into the tissue culture medium is analyzed using a commercially available radioimmunoassay. Although this assay is simple to carry out, it should be used with caution, since it is based on the assumption that the appearance of secreted protein in the medium parallels the accumulation of a cytoplasmic enzyme. Thus, the assay does not necessarily provide an accurate internal control for the measurement of enzymatic activity in cell extracts.

EXPRESSION OF PROKARYOTIC GENES: PROMOTERS

The first step in expressing eukaryotic proteins in bacteria is to choose an expression vector that carries a strong, regulated prokaryotic promoter. Here we describe the use of expression vectors that contain a bacteriophage λ p_L promoter, a hybrid *trp-lac* promoter, or a bacteriophage T7 promoter.

The Bacteriophage λ p_L Promoter

The bacteriophage λ p_L promoter is regulated by a temperature-sensitive repressor, *cIts*857, which represses p_L transcription at low temperatures but not at elevated temperatures. *E. coli* strain M5219 contains a defective bacteriophage λ prophage, which encodes the *cIts*857 repressor, and the bacteriophage λ N protein, an antagonist of transcription termination. This system is particularly suitable if the product of the gene to be expressed is toxic to *E. coli*, since *cIts*857 strongly represses transcription. In addition, the antitermination function of the *N* gene may allow RNA polymerase to read through potential termination sites within the gene. One drawback of this system is that the temperature shift (step 5) induces not only the p_L promoter, but also the heat-shock genes, some of which encode proteases (Buell et al. 1985). This problem can be alleviated by using a bacteriophage λ cI^+ lysogen and inducing with mitomycin C or nalidixic acid (Shatzman and Rosenberg 1987). pKC30 (see Figure 17.4) (Shimatake and Rosenberg 1981) is one of many p_L vectors available.

To provide a p_L promoter to direct transcription of a cloned gene that has an *E. coli* ribosome-binding site:

1. Digest pKC30 with *Hpa*I.
2. Digest the cloned DNA with appropriate restriction enzymes at a position 5′ of the initiation codon and at a site 3′ of the terminus of the cloned gene.
3. Insert the cDNA into the plasmid, ligate the DNA, and transform *E. coli* strain M5219. Plate transformants on LB medium containing ampicillin (100 μg/ml) and incubate overnight at 30°C.
4. Screen transformants for the presence of the desired insert by colony hybridization and/or by plasmid minipreps and restriction enzyme analysis.
5. To obtain high levels of transcription of the cloned gene, grow strain M5219 containing the expression plasmid to mid-log phase at 30°C, and then shift the temperature of the culture to 40°C. Continue to incubate for several hours at 40°C. Remove small aliquots at various times and analyze them by one of the methods discussed on pages 17.34–17.35. The kinetics of induction varies with different proteins, so it is necessary to determine the time at which the maximum amount of product is present.

Although 42–45°C is used elsewhere in this manual to inactivate bacteriophage λ *cIts*857, 40°C is used here in order to reduce the induction of heat-shock proteins and to allow the cells to continue growth.

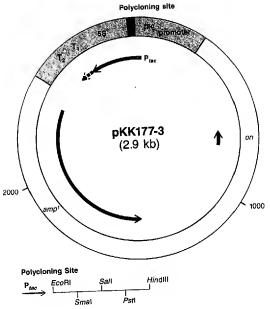

FIGURE 17.5
pKK177-3 is a *tac* vector containing multiple sites downstream from the *tac* promoter into which a gene can be cloned. Downstream from these sites is *rrn*B, which contains an *E. coli* 5S gene and the T_1 and T_2 terminators (Amann and Brosius 1985).

The trp-lac Promoter

Another promoter that has been used successfully to produce large amounts of proteins in *E. coli* is the *tac* promoter, a hybrid *trp-lac* promoter that is regulated by *lac* repressor (Amann et al. 1983; de Boer et al. 1983). Transcription is repressed in *E. coli* strains such as RB791, a *lacI*^{q} strain that makes high levels of *lac* repressor. If the protein to be expressed is toxic to *E. coli*, then a *lacI*^{q} gene should be cloned into the expression plasmid to make higher levels of *lac* repressor. One factor to consider in choosing this vector system is that the *tac* promoter is induced by adding isopropylthio-β-D-galactoside (IPTG), a relatively expensive compound. A useful *tac* promoter expression plasmid, pKK177-3 (Amann and Brosius 1985), is illustrated in Figure 17.5.

The steps used in cloning and expression in *tac* vectors (or *lac* vectors) are similar to those for other promoter systems and are described below.

1. Clone the cDNA to be expressed, along with its ribosome-binding site, into the polycloning site of pKK177-3.

2. Transform *E. coli* strain RB791, and screen transformants by colony hybridization and/or plasmid minipreps and restriction enzyme analysis.

3. To obtain high levels of transcription, grow cells to mid-log phase at 37°C, add IPTG to a final concentration of 1 mM, continue growth, and monitor the amount of protein made at various times after induction using one of the methods discussed on pages 17.34–17.35.

   For preparation of a stock solution of IPTG, see Appendix B.

Note

Hasan and Szybalski (1987) have designed *tac* vectors in which expression of a cloned gene is controlled by promoter inversion in vivo. The promoter in the vector is directed away from the gene to be expressed so that transcription of the gene is kept extremely low until production is desired. The promoter is flanked by *att*P and *att*B and can be inverted efficiently by the bacteriophage λ Int protein. These vectors also utilize bacteriophage λ *N*-mediated antitermination as described on page 17.11 for pKC30.

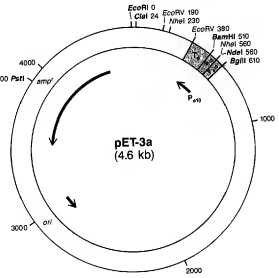

FIGURE 17.6
pET-3 carries the bacteriophage T7 ϕ10 promoter ($P_{\phi10}$) and ϕ terminator (T_ϕ). The terminator may make the transcripts more resistant to exonucleolytic degradation (Studier and Moffatt 1986). pET-3a is a derivative of pET-3 into which the translation start (S_{10}) of bacteriophage T7 ϕ10 (the major capsid protein of bacteriophage T7) with a *Bam*HI site at codon 11 has been inserted. The *Nde*I site (CATATG) is located at the translation start site and can be used to construct a plasmid that directs the expression of native proteins.

The Bacteriophage T7 Promoter

A novel expression system has been developed by Studier and Moffatt (1986) and Tabor and Richardson (1985) using a bacteriophage T7 RNA polymerase/promoter system. This system is designed for the exclusive expression of the cloned gene. Bacteriophage T7 RNA polymerase recognizes solely bacteriophage T7 promoters, will transcribe around a plasmid several times, and may transcribe sequences that are not efficiently transcribed by *E. coli* RNA polymerase. This system allows high levels of expression of some genes that are not expressed efficiently in other systems.

Two components are required for the bacteriophage T7 expression system:

- The first component is bacteriophage T7 RNA polymerase. This polymerase is the product of bacteriophage T7 gene *1* and can be provided on an infecting bacteriophage λ vector or produced from a gene copy inserted into the *E. coli* chromosome (Tabor and Richardson 1985; Studier and Moffatt 1986). If expression of the cloned gene product is toxic, then the level of bacteriophage T7 RNA polymerase must be kept low during cell growth. One way to accomplish this is to use the lysogen BL21(DE3), in which bacteriophage T7 gene *1* is expressed from a *lac*UV5 promoter. In some cases, it is necessary to use cells in which no bacteriophage T7 RNA polymerase is present until expression is desired. This is accomplished by infecting the host cells (e.g., HMS174) that carry the expression plasmid with bacteriophage CE6 (λ*c*I*ts*857 *S*am7) carrying bacteriophage T7 gene *1* (Studier and Moffatt 1986).
- The second component is a plasmid vector with a bacteriophage T7 promoter upstream of the gene to be expressed. pET-3 is a derivative of pBR322 that carries the bacteriophage T7 gene *10* promoter ($P_{\phi 10}$), a *Bam*HI cloning site, and the bacteriophage T7 transcription terminator (T_ϕ) (Rosenberg et al. 1987). Derivatives of pET-3 (e.g., pET-3a; see Figure 17.6) have been constructed to provide a bacteriophage T7 gene *10* translation start (S_{10}) through codon 11. DNA can be inserted in each of the three reading frames at this codon to express fusion proteins. The pET-3 derivatives can also be used for production of intact native proteins as described on pages 17.17–17.24 by fusing coding sequences at the *Nde*I site immediately preceding the ATG. Another bacteriophage T7 expression plasmid, pT7-1, has a polycloning site downstream from the bacteriophage T7 ϕ10 promoter (Tabor and Richardson 1985).

To express a gene in *E. coli* using bacteriophage T7 RNA polymerase:

1. Clone the gene of interest into a bacteriophage T7 promoter expression plasmid. Identify the correct plasmids in a standard *E. coli* strain by miniprep analysis.

2. Transform *E. coli* strain BL21(DE3), and select for ampicillin-resistant transformants. BL21(DE3) is a lysogen bearing the bacteriophage T7 polymerase gene under the control of the *lac*UV5 promoter. (See precautions recommended by Studier and Moffatt [1986].)

3. Inoculate NZCYM medium with one colony, and incubate overnight at 37°C to obtain a saturated culture.

4. Induce the culture and determine the amount of protein produced as follows:

   a. Inoculate 5 ml of NZCYM medium containing ampicillin (100 μg/ml) with 50 μl of saturated culture. Incubate the culture for 2 hours at 37°C.

   b. Remove 1 ml of the uninduced culture and place in a microfuge tube. Process as described in steps c and d below. Induce the remaining culture by adding isopropylthio-β-D-galactoside (IPTG) to a final concentration of 1 mM.

      For preparation of a stock solution of IPTG, see Appendix B.

   c. Remove 1-ml aliquots of the induced culture at 0.5, 1, 2, and 3 hours after induction. Immediately centrifuge these samples at 12,000*g* for 1 minute at room temperature in a microfuge. Remove the supernatants by aspiration.

   d. Resuspend each pellet in 100 μl of 1 × SDS gel-loading buffer, heat to 100°C for 3 minutes, and store at 0°C until all of the samples are collected and ready to load on a gel.

      1 × SDS gel-loading buffer

      50 mM Tris · Cl (pH 6.8)
      100 mM dithiothreitol
      2% SDS (electrophoresis grade)
      0.1% bromophenol blue
      10% glycerol

      1× SDS gel-loading buffer lacking dithiothreitol can be stored at room temperature. Dithiothreitol should then be added just before the buffer is used, from a 1 M stock (see Appendix B).

   e. Thaw the samples, and then centrifuge them at 12,000*g* for 1 minute at room temperature. Load 15 μl of each suspension on an SDS-polyacrylamide gel of the appropriate concentration (see Chapter 18, pages 18.47–18.54). Use as a control a suspension of cells containing the vector alone.

5. Stain the gel with Coomassie Brilliant Blue or silver stain, or carry out a western blot (if antibody already exists), to identify induced protein (see Chapter 18, pages 18.55–18.57 or 18.60–18.75, respectively).

Note

For expression of some proteins, it is advisable to increase the ampicillin concentration to levels as high as 200 μg/ml to select for cells that retain the expression plasmid.